Fractal-Based Video Coding and Slow Motion Replay*

Jean-Luc Dugelay, E. Majdandzic and S. Roche
Dept. of Multimedia Communications, Institut EURECOM,
2229, Route des Crêtes, B.P. 193, 06904 Sophia Antipolis – France
dugelay@eurecom.fr; http://www.eurecom.fr/~image

1. Introduction

Since the landmark paper on image coding with iterated function systems (IFS), in 1992 by A. Jacquin [1] many authors have studied IFS (or "fractal") and proposed many improvements to Jacquin's algorithm [2]. More recently, in addition to compression, the fractal coding technique has also been investigated for developing multimedia imaging tools such as zoom [3], security [4], retrieval and indexing [5]. In this paper, we (1) present an extension from still images of the basic algorithm to moving picture, (2) use fractal properties in order to introduce the possibility of a temporal zooming (i.e. slow motion replay) of the video at the decoding stage, and (3) compare early results obtained in this way with a classical two steps scheme composed of first applying MPEG-1 for the compression stage, and second, basic frame-interpolation procedure for doing the temporal resolution enhancement stage of video.

1.1 Review of fractal image coding

Basic concepts. Let, μ_{orig} be the image to be compressed, μ_0 an arbitrary initial image, μ and ν two generic images, and $d(\mu,\nu)$, a distortion measure which measures the dissimilarity between images. A transformation τ that maps an image into another is said to be *contractive* if: $d(\tau(\mu),\tau(\nu)) < \sigma . d(\mu,\nu)$ with $0 < \sigma < 1$, where σ is the *contractivity* of τ. Then $\tau^n(\mu_0)$ converges to an *attractor* μ_a as n approaches infinity, where μ_a is independent of μ_0. The Collage Theorem states that if there exists a transformation τ such that, $d(\mu_{orig},\tau(\mu_{orig})) < \varepsilon$ and τ is contractive with contractivity σ, then $d(\mu_{orig},\mu_a) < \varepsilon/(1-\sigma)$.
The task of the encoder is to determine a transformation τ (an "IFS code") for which $\tau(\mu_{orig})$ is as similar as possible to μ_{orig} subject to a limitation on the number of bits needed to specify τ. The IFS code τ is transmitted to the decoder which then computes the attractor μ_a as the reconstructed image. The reconstruction error is upper-bounded by the Collage Theorem.

Encoding stage. In Jacquin's approach, the encoder finds the transformation τ from the original image μ_{orig} as a sum of affine transformations (τ_i), one for each range block, R_i, each of which maps a particular domain block into the corresponding range block R_i. The domain and range are partitioned at different resolutions; typically with square range blocks of size *B*x*B* and domain blocks of size *2B*x*2B* (Generally *B* is 8 pixels). For each range block, the encoder searches for the best collage match from a suitably transformed and selected domain block. For this search, candidate domain blocks are transformed in three steps, by performing sub-sampling, isometry, and scale and shift operations on the block luminance values, **see figure no. 1.**

Decoding Stage. The decoding stage is based on a continuous theory. To decode an image, the received IFS code, τ, is applied to an arbitrary initial image μ_0 to form an image μ_1. The process is repeated to obtain μ_2 from μ_1, and so on, until it reaches μ_a. Typically, less than ten iterations are needed for convergence.

---

* This work was supported in part by AEROSPATIALE, CNET, and DGA.

1.2 Zooming using I.F.S.

In addition to compression, IFS possess some properties of fractals which can be used to sub/over-sample an image. In this section, we review how I.F.S. can be used for image zooming directly from basic algorithms of still image compression.
The main idea of the fractal zoom is rather simple, see **figure no. 2**. It is based on an important property of the fractals. If we assume that the fractal coding is really a fractal process, then the fractal code's attractor is a fractal object. In fact, by iterating a determinist transformation on some initial image, a determinist fractal image is obtained. Therefore, this process must be independent of the resolution required. The original image μ_{orig} has a fixed size defined by the number of its pixels. But the fractal code has no intrinsic size, because it is a transformation which can be theoretically applied on any image. Hence, we can assume that:

- the coding error is rather small;
- self-similarities are scale-independent.

Then, the fractal code enables to zoom. In practice, this operation consists of increasing the range block's size, and therefore the domain block's size. For a zoom of factor z, the new sizes for the decoding stage will be obtained by multiplying by z the initial sizes (i.e. used during the encoding stage); but the fractal code τ will remain unchanged.

2. From image to video

2.1 Introduction

Lately, two approaches have been proposed in the literature [6] for a possible extension of Jacquin's algorithm for moving picture coding:

Block coding schemes.

i) realized picture by picture, eventually using motion information in order to limit the search for the matching between range and domain blocks.

ii) an "intra" coding mode using I.F.S. associated with an "inter" coding mode using block-matching. This method is close to MPEG standard using D.C.T. (Discrete Cosine Transform) and the block-matching technique.

Cube coding scheme. The sequence to be coded is partitioned into several GoPs (Groups of Picture). Instead of segmenting an image into a set of square blocks, each GoP is segmented into a set of cubes. The third dimension is associated to the temporal dimension.

The second approach "cube approach" is more complex to implement than the first one (i.e: mixed approach, IFS plus Block-matching). Nevertheless, only the "cube approach", despite the relatively high computational cost (inherent to I.F.S. based coding methods), retains the most interesting properties of I.F.S. based coding, included in fractal theory. In particular, the code we obtained is (theoretically) resolution independent, and so we can perform a temporal sub/over-sampling (at the decoding stage).

2.2 Cube approach

The complete video coding scheme has already been published and discussed in [7]: Coding image sequences with 3-D I.F.S. can be considered as a direct extension of Jacquin's 2-D scheme. In fact, the image sequence is subdivided into GoPs and each one is coded as a still image, while replacing 2-D blocks (squares) by 3-D ones (cubes), see **figure no. 3.** A 3-D block is made up of 2-D blocks, belonging to successive frames, and having the same spatial position.

2.3 Temporal zooming

There have been many publications on fractal image coding, some on fractal zoom and on video coding, but there have not been any published studies on the temporal resolution enhancement feature of the fractal coders. In this section we propose a fractal approach which unifies compression and possibility of slow motion replay at the decoding stage. Similar to image zooming, the compressed data τ will be unchanged in-between the coding and decoding stages; only the size of blocks will be increased according to the temporal dimension. We then compare this approach with a classical scheme composed of a standard encoder (i.e. MPEG-1) and a standard frame-interpolation procedure (i.e. duplication, repetition), see **figure no. 4**. Early results are discussed in the next section.

3. Simulations and early results

In order to perform simulations with an acceptable computational time , we had to simplify the theoretical 3-D IFS algorithm presented in section 2.2, with no additional degradations of the decoded video. These simplifications consist of:
- canceling isometries;
- limiting search of the domain block in the neighborhood of the range cube;
- simplifying luminance parameters estimation.

Comparative tests have been performed on sequences "Salesman" and "Miss America". Several sets of parameters have been defined in order to obtain approximatively the same compression rate when using MPEG-1 and I.F.S. For the fractal approach, we used a size of range (respectively domain) cubes of size 4x4x4 (resp. 8x8x8) for the encoding stage and then a size of range (resp. domain) cube of size 4x4x4z (resp. 8x8x8z) for the decoding/zooming stage, where z is the zoom factor, for the decoding stage (z=1 consists of decoding the video without zoom). The early results we obtained show that fractal zoom (z=2 and z=4) leads to some annoying temporal artifacts between GoPs, even if locally, for some objects, slow replay motion is visually better reconstructed than using a frame duplication approach on decompressed MPEG stream. This problem of temporal links between GoPs is not so important when using MPEG approach because, as shown in figure no. 4, extreme frames are not interpolated; these frames are "original" decoded I-frames, and therefore present little visual degradations. This allows a smooth transition between frames of two consecutive GoP. Using IFS, encoding and then decoding quality of all frames in a GoP is the same. Moreover each GoP is processed independently. Possible improvements (not yet simulate) could consist of (1) modifying the error criterion used to determinate the best matching between domain and range blocks: we could allow a larger weight for the extreme frames of each GoP in order to reduce the "cut effect" between GoP when decoding/zooming; or (2) creating an overlap between GoPs before encoding.

4. Concluding remarks

Using a fractal approach for video encoding, we obtain a compressed data stream which possesses the property of being decodable, without any additional information or post-processing step, at any size according to the spatial dimensions as well as according to the temporal one. Hence, this approach can be used to realize a replay, and more generally any video format conversion (in terms of frame rate and/or image size) allowing the possibility of an automatic adaptation of the video stream to the terminal of visualization. Nevertheless, this study also shows that results obtained in such a way (i.e. using a basic extension of fractal coding) do not produce better performance than those obtained by using a classical framework in terms of image qualities. In some other words, this study highlights the difficulty to extend fractal coding technique of images to video (without losing fractal properties) by considering the temporal dimension as an additional one.

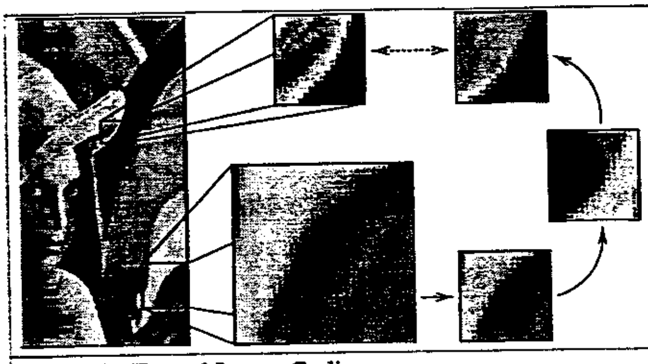

Figure 1 : Fractal Image Coding

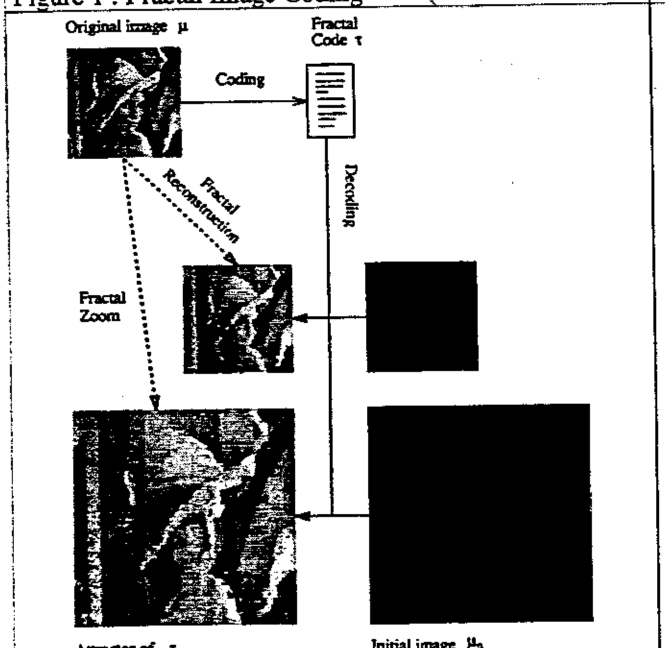

Figure 2 : Fractal Image Zooming

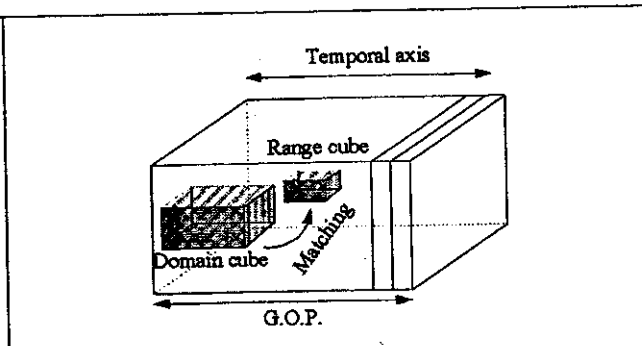

Figure 3 : Fractal Video Coding

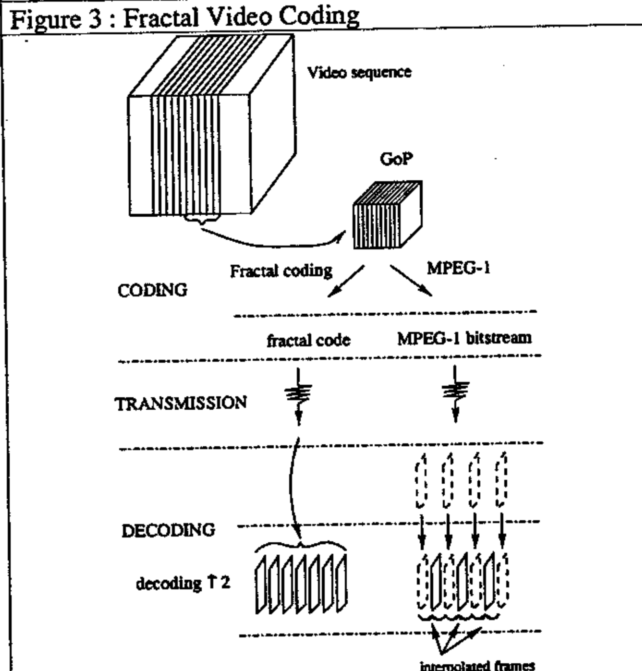

Figure 4 : Fractal/MPEG Video Coding and Zooming

5. References

1. A.E. Jacquin, Image Coding based on a Fractal Theory of Iterated Contractive Image Transformations, *IEEE Trans. on Image Proc.*, 2,1 (1992), pp. 18-30.
2. Y. Fisher (Ed.), Fractal Image Compression, Springer Verlag, NY, 1994.
3. E. Polidori and J.-L. Dugelay, Zooming using I.F.S., Fractals, Vol. 5, April 1997, pp. 111-123.
4. P. Davern and M. Scott, Fractal Based Image Steganography, *Information Hiding, 1st Int. Workshop*, Cambridge, May/June 1996.
5. J.K. Wu et al., A Fractal-Based Clustering Approach in Large Visual Database Systems, *Multimedia Tools and Applications* 3, pp. 225-244 (1996).
6. J.-L Dugelay and E. Majdandzic, Moving Picture Fractal Compression using I.F.S. – A Review (II), May 1997, *The Freiburg Paper Collection on Fractal Image Compression*, ftp://ftp.informatik.uni-freiburg.de/documents/papers/fractal.
7. M. Barakat and J.-L. Dugelay, Image Sequence Coding using 3-D I.F.S., *IEEE ICIP'96*.

Acknowledgments. The authors thank M. Barakat, E. Polidori and J.-M. Sadoul for their respective contributions to this study.